chinaXiv:201705.00900v1

国内情报学相关性研究现状浅析

——基于社会网络分析方法的定量分析

Analysis on the Research Status of Domestic Relevance of Information Science

——Based on the Quantitative Analysisof the Social Network Analysis Method

和婧，中国科学院文献情报中心，北京中关村北四环西路33号，100190，hejing@mail.las.ac.cn

摘要：本文基于社会网络分析方法，借助 SATI、Ucinet、NetDraw 和 NoteExpress 统计分析软件，通过分析我国情报学相关性领域论文成果的关键词共现和作者合作情况，认为该领域目前的研究热点为信息检索的相关性研究、引文分析中的相关性研究、数据挖掘与知识管理中的相关性研究以及与国外情报学相关性研究的对比研究。针对该领域存在的研究重点单一、学者关注度不够等问题，文章中还提出开展以“情报学相关性研究”为主题的研讨会，鼓励跨学科、跨地区学者交流等建议，以期为我国情报学相关性领域的发展提供参考和借鉴。

关键词：情报学　相关性　社会网络　关键词共现　作者合作

Abstract：The article analyses the key words co-occurrence and cooperation of the authors of the articles in the domestic relevance of Information Science which is based on the social network analysis methods and by means of the software of SATI、Ucinet、NetDraw and NoteExpress. The article analyses the research hotpots which are the relevance studies of Information retrieval, citation analysis, data mining and knowledge management, and the foreign relevance study of Information Science. It points out the dificiencies which arenarrowness of the research area and the insufficient attention of the researchers. The advice of conducting the relevance seminar, encouraging the interdisiplinary and interregional communication is put up to give reference to the development of the domestic relevance of Information Science.

Keywords：Information Science; relevance; social network; key words co-occurrence; cooperation of the authors

chinaXiv:201705.00900v1

1. 引言

相关性是情报学产生的重要原因[1]，正是由于信息之间所具有的相关关系，才使得情报学得以实现借助相关的分析工具，对情报学相关性的研究最早起源于国外，其发展可分为三个阶段，分别是 1958 年之前，1959 年—1976 年，以及 1977 年至今[2]，对情报学相关性的研究高峰分别集中在 20 世纪 60 年代—70 年代前期，20 世纪 80 年代后期至今这两个时期[3]。国外对该领域的研究，对情报学相关性理论的发展起到了很大的推动和引领性作用，我国对该领域的研究也深受其影响。本文旨在通过对情报学相关性领域现有的期刊论文进行分析，以探析我国国内在该领域的研究热点、代表性人物、存在的问题等，进而提出相应建议，以期为该领域的发展提供参考。

社会网络分析方法（Social Network Analysis）简称 SNA，是对社会网络中各种关系结构及其属性加以分析的一套理论与方法[4]，与社会网络分析方法相对应有专业的统计分析工具，如 SATI、Ucinet 和 NetDraw，这些方法和工具可用来分析关键词共现和作者合作的情况。关键词能够反映出某一领域的研究热点，作者合作情况能够反映出某一领域的发展状况。因此，本文借助社会网络分析方法及相应的分析软件，对情报学相关性研究中的关键词共现和作者合作进行分析，以探究我国在情报学相关性研究的现状。

2. 数据来源与分析方法

本文所使用的数据全部来源于中国期刊全文数据库（CNKI），选定的来源数据库为中国学术期刊网络出版总库，检索策略以主题和关键词为检索点，以“相关性”为检索词。由于本文探讨情报学领域对相关性的研究，又因核心期刊的文章更具权威性和代表性，因此将来源期刊限定在图书情报学领域的 19 种中文核心期刊，这一点可通过 CNKI 提供的高级检索界面中的来源期刊进行设置。经过上述检索条件的限定，共得到 252 条记录，经过一一比对，发现检索结果中仍存在与情报学相关性无关的文章，经过筛选，共有 183 篇有效论文。

本文在分析数据的过程中，主要使用了 SATI、Ucinet 和 NetDraw 三种分析工具，由于 SATI 不能完整统计数据，只能统计出现频次排名在前 100 名的数据，因此辅助使用了 NoteExpress 软件，用于关键词总数和作者总数的统计。使用的社会网络分析方法主要涉及社会网络的密度分析、中心度分析、凝聚子群分析。使用的软件功能主要有 K 核分析、度数中心度分析、中介中心度分析、接近中心度分析、K 从分析等。

结合分析工具和分析方法，具体实施步骤如下，首先从 CNKI 中下载数据，

格式保存为软件可读的 Endnote 格式，将下载数据导入到 SATI 中，根据需要抽取关键词、作者、地址（软件中的地址字段对应的是作者所在机构）等字段，进行频次统计之后生成共现矩阵；将共现矩阵的数据导入到 UCINET 中，将数据格式转化为##h 的格式；最后导入到 NetDraw 中，进行可视化呈现与分析。

3. 研究现状分析

3.1 研究热点分析

关键词是作者对其文章研究定位的重要体现，也是对某领域研究点研究意义的肯定。在关键词共现网络中，被指向次数越多的关键词，其在网络中的重要性越大，该关键词即为目前该领域的热点研究方向[5]。因此，对所有情报学相关性的期刊论文关键词进行分析，可得到情报学相关性研究的热点。

通过使用软件 SATI 形成关键词共现矩阵，如表 1 所示，将矩阵导入到 Ucinet 软件中，并采用 Netdraw 生成关键词共现的可视化知识图谱。Netdraw 中有 K-core 分析功能，最终得到图 1 的可视化结果。利用 K-core 分析功能将这些关键词进行自动归类，将网络中的重点关键词标红显示，且在图中较为突出，由此可确定网络中的重点关键词。由图可以看出，突出显示的核心关键词有相关性、情报学、信息检索、检索效率、查全率、查准率、搜索引擎、数据挖掘、知识管理等。

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		相关性	信息检索	搜索引擎	查全率	查准率	H指数	informati	情报检索	Internet	引文分析	情报检索系	查询扩展	情报学	情报科学	相似度	特征因子	主题爬行器
2	相关性	1	0.1562	0.0078	0.1225	0.1029	0	0	0.0533	0	0	0.1	0.036	0.036	0.08	0	0	0
3	信息检索	0.1562	1	0.0106	0	0	0	0	0	0.0216	0	0	0.0486	0.0216	0	0.0068	0	0
4	搜索引擎	0.0078	0.0106	1	0	0	0	0	0	0.0087	0	0	0	0	0	0.0109	0	0
5	查全率	0.1225	0	0	1	0.875	0	0	0.0208	0	0	0.225	0	0	0.125	0	0	0
6	查准率	0.1029	0	0	0.875	1	0	0	0.0238	0	0	0.2571	0	0	0.1429	0	0	0
7	H指数	0	0	0	0	0	1	0	0	0	0.1143	0	0	0	0	0	0.0357	0
8	informati	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
9	情报检索	0.0533	0	0	0.0208	0.0238	0	0	1	0	0	0	0.0333	0	0.0417	0	0	0
10	Internet	0	0.0216	0.0087	0	0	0	0	0	1	0	0	0	0	0	0	0	0
11	引文分析	0	0	0	0	0	0.1143	0	0	0	1	0	0	0	0	0	0	0
12	情报检索系	0.1	0	0	0.225	0.2571	0	0	0	0	0	1	0	0	0.45	0	0	0
13	查询扩展	0.036	0.0486	0	0	0	0	0	0.0333	0	0	0	1	0	0	0	0	0
14	情报学	0.036	0.0216	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
15	情报科学	0.08	0	0	0.125	0.1429	0	0	0.0417	0	0	0.45	0	0	1	0	0	0
16	相似度	0	0.0068	0.0109	0	0	0	0	0	0	0	0	0	0	0	1	0	0
17	特征因子	0	0	0	0	0	0.0357	0	0	0	0	0	0	0	0	0	1	0.25
18	主题爬行器	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.25	1
19	评价	0.005	0	0.0109	0.0312	0.0357	0	0	0	0	0	0.05	0	0	0	0	0	0
20	关联规则	0.005	0.0068	0	0	0	0	0	0.0417	0	0	0	0.2	0	0	0	0	0
21	综述	0.0067	0	0.058	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	analysis	0	0	0	0	0	0	0.0556	0	0	0	0	0	0	0	0	0	0
23	相关度	0.0067	0.009	0.0145	0	0	0	0	0	0	0	0	0	0	0	0.0833	0	0
24	本体	0.0267	0.009	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	网络信息ネ	0	0	0.058	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	元搜索引擎	0.0267	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	评价指标	0	0	0.1304	0	0	0	0	0	0	0	0	0	0	0	0	0	0

表 1.关键词共现矩阵

chinaXiv:201705.00900v1

chinaXiv:201705.00900v1

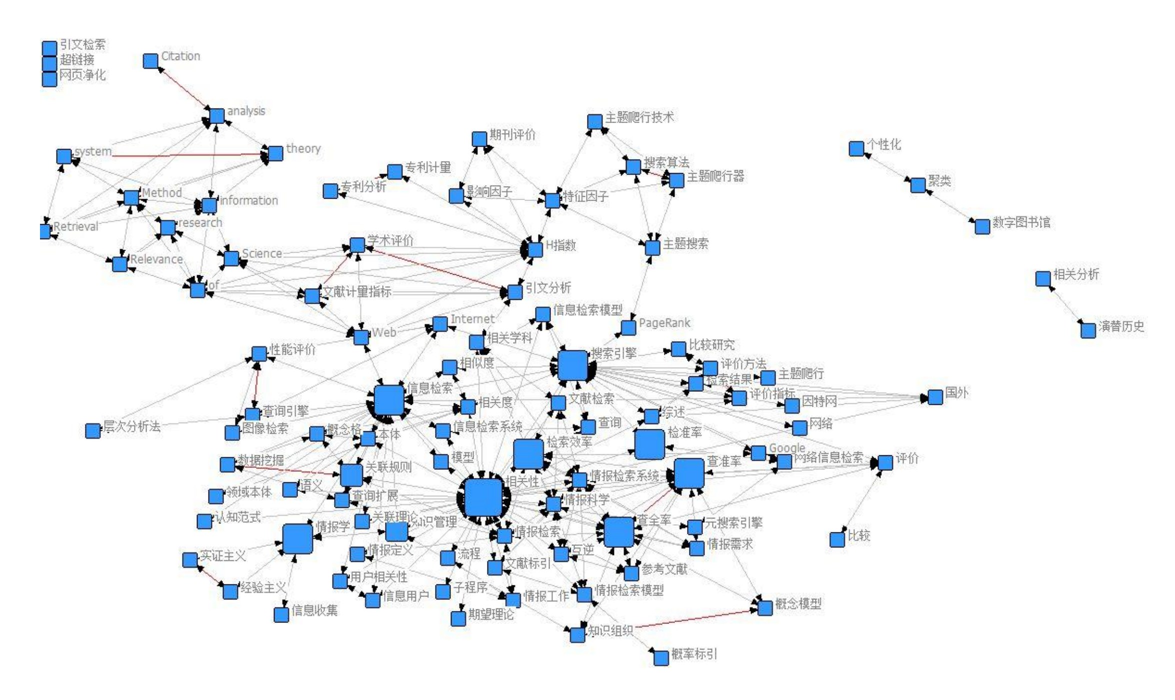


图 1 关键词共现可视化网络

观察关键词共现网络发现，数据结点有两方面需要处理，一方面是孤立存在的数据结点，由于孤立的数据结点与整体网络没有联系，没有研究意义，因此删掉孤立结点；另一方面是存在相同或相近关键词，如情报学与 Information Science，检索与 Retrieval 是相同关键词，信息检索、文献检索与信息检索系统，评价与评价指标、评价方法等是相近关键词。然而由于 SATI 和 Ucinet 软件均不具有合并相近关键词的功能，因此有些关键词显示不够突出。为了突出显示情报学相关性领域内的研究热点，根据专业知识将网络中节点的位置进行调整，得到图 2。

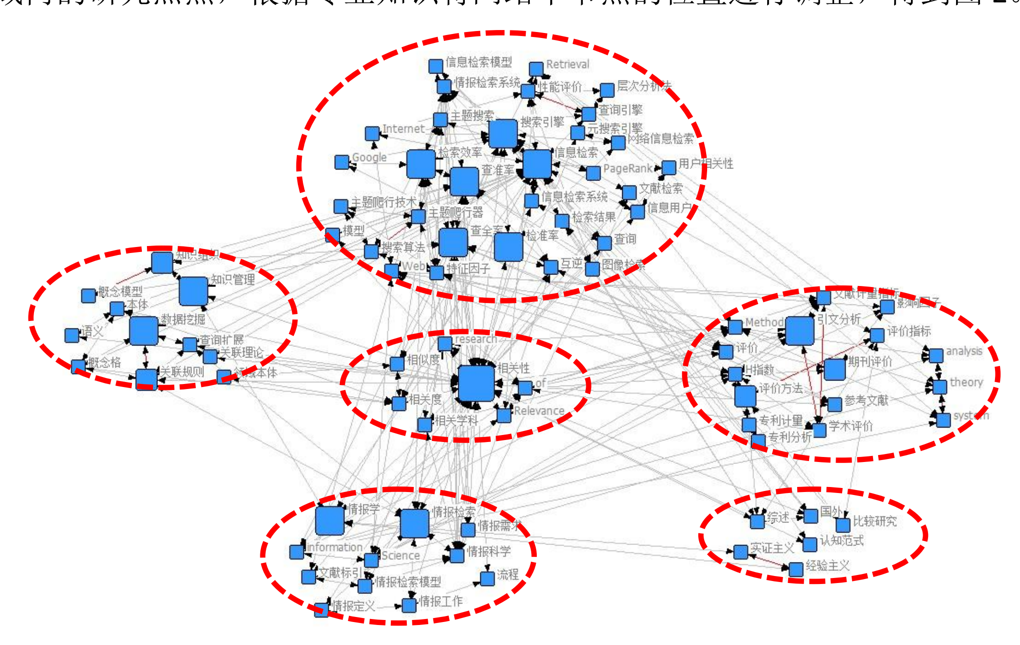


图 2 关键词共线网络聚类图

经过聚类之后，共出现 6 个大类，其中相关性与情报学这两个大类没有实质性意义，对研究该领域热点无关，故应去掉。最终可得出，我国国内情报学相关性领域四个主要的研究热点，分别是信息检索中相关性研究、引文分析中相关性研究、数据挖掘与知识管理中相关性研究以及国内外情报学相关性的比较研究。

首先，关于信息检索中相关性研究。在已检索到的 183 篇文章中进行二次检索，限定主题为“检索”或“搜索”，得到 93 条记录，约占总论文的 50.8%，所占比重较大。因此，情报学相关性领域内的热点之一即信息检索中相关性研究。将 93 篇文章的数据下载后导入到 NoteExpress 中统计，共得到 387 个关键词，高频关键词见表 2。其中除了信息检索和相关性之外，占比重较大的有搜索引擎、Internet、本体、查全率、模型网络信息检索、评价指标、搜索算法、文本挖掘等。由此可知，信息检索相关性研究的主要内容有信息检索相关性模型研究、网络搜索引擎与信息检索系统评价研究、相关性算法改进研究、关联规则在信息检索相关性中的应用以及通过信息检索相关性实现文本信息挖掘。

chinaXiv:201705.00900v1

字段: 关键词	记录数	% (387)	图形
信息检索	30	7.752 %	
相关性	22	5.685 %	
搜索引擎	19	4.910 %	
Internet	4	1.034 %	
本体	3	0.775 %	
查全率	3	0.775 %	
模型	3	0.775 %	
网络信息检索	3	0.775 %	
评价指标	3	0.775 %	
Google	2	0.517 %	
PageRank	2	0.517 %	
WWW	2	0.517 %	
个性化	2	0.517 %	
主题搜索	2	0.517 %	
主题爬行器	2	0.517 %	
信息检索模型	2	0.517 %	
元搜索引擎	2	0.517 %	
关联理论	2	0.517 %	
分布式信息检索	2	0.517 %	
国外	2	0.517 %	
引文检索	2	0.517 %	
情报检索	2	0.517 %	
搜索算法	2	0.517 %	
文本挖掘	2	0.517 %	
查准率	2	0.517 %	
查询扩展	2	0.517 %	
检索结果	2	0.517 %	
特征因子	2	0.517 %	
用户相关性	2	0.517 %	
相似度	2	0.517 %	
相关度	2	0.517 %	
综述	2	0.517 %	
评价	2	0.517 %	

表 2 信息检索相关性论文关键词统计表

其次，关于引文分析中相关性研究。引文分析可作为评价某一学科领域内重要学者、重要期刊以及论文研究成果学术影响力的重要依据。引文分析是相关性在情报分析中的重要应用，经统计发现，其研究内容主要涉及引文评价方法、引文索引的检索效率、引文分析中的文献计量指标及算法、引文分析的应用实例等。

再次，关于数据挖掘与知识管理中的相关性研究。相关性在情报学中的另一应用就是数据挖掘与知识管理方面的应用。数据挖掘与知识管理都是从潜在的、隐含的信息中，经过相应的分析过程，挖掘信息之间隐含的相关关系，进而分析得出隐藏的有价值的信息内容[6]。目前的论文成果主要涉及数据挖掘与知识管理

chinaXiv:201705.00900v1

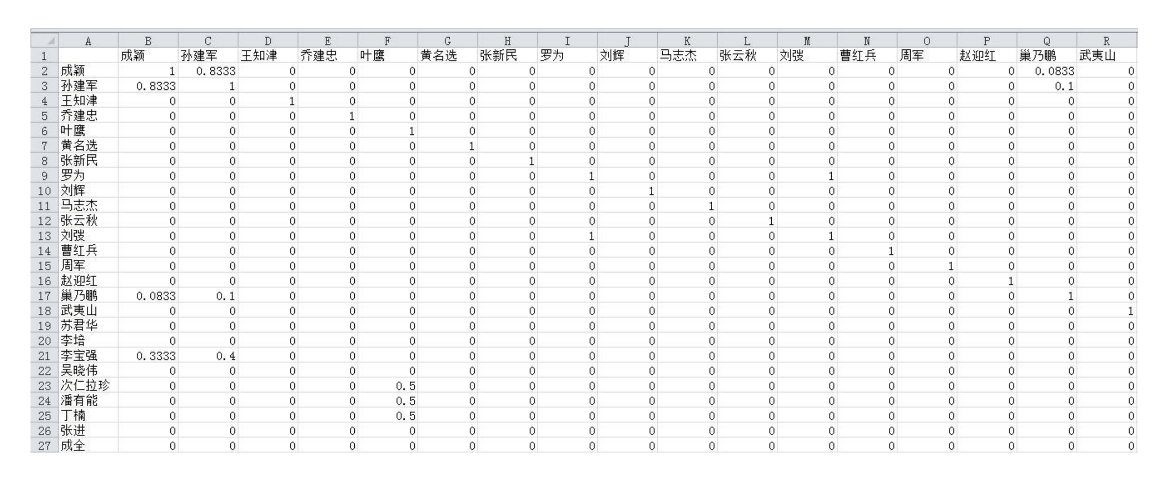
的方法与类型，其中用于网页信息挖掘的 Pagerank 方法、HITS 算法、SALSA 等方法与信息系统检索性能评价的内容相关。

最后，关于国内外情报学相关性的比较研究。根据文献调研，国内学者较为关注国外相关性理论的发展。由于国外在情报学相关性领域的研究较早，理论成果较为丰富和成熟，因而国内学者较为注重对国外研究的总结，出现了一些综述类的研究成果。

3.2 研究学者及其所在机构分析

关注某一领域的学者数量、学者在某一领域发表论文的数量、学者在某一领域的合作情况等在一定程度上反映了该学科的发展情况。若该领域是热点研究领域，则学者该领域的关注度较高，发文较多，文章的被引频次较高，并且发表有关该领域论文的学者数量较多。

通过使用软件 SATI 形成高频作者共现矩阵，如表 3 所示，将矩阵导入到 Ucinet 软件中，并采用 Netdraw 生成高频作者合著网络图谱。使用 Netdraw 中的 K-core 分析功能，最终得到图 3 的可视化结果。图中的节点代表学者，节点的大小代表在该网络中的中心度，节点间的连线粗细表示作者间合作的关系强弱。K-core 分析功能将作者进行自动归类，其中重点作者标红且突出显示，这表明这些作者在该领域具有较大的影响力。需要说明的是，由于分析作者的合作情况，因此单独发文的作者节点没有分析意义，已从图中删除。



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1		成颖	孙建军	王知津	乔建忠	叶鹰	黄名选	张新民	罗为	刘辉	马志杰	张云秋	刘弢	曹红兵	周军	赵迎红	巢乃鹏	武夷山
2	成颖	1	0.8333	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0833	0
3	孙建军	0.8333	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0
4	王知津	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	乔建忠	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	叶鹰	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
7	黄名选	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
8	张新民	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
9	罗为	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
10	刘辉	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
11	马志杰	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
12	张云秋	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
13	刘弢	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
14	曹红兵	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
15	周军	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
16	赵迎红	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
17	巢乃鹏	0.0833	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
18	武夷山	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
19	苏君华	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	李培	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	李宝强	0.3333	0.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	吴晓伟	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	次仁拉珍	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
24	潘有能	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
25	丁楠	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
26	张进	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	成全	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

表 3 高频作者共现矩阵

chinaXiv:201705.00900v1

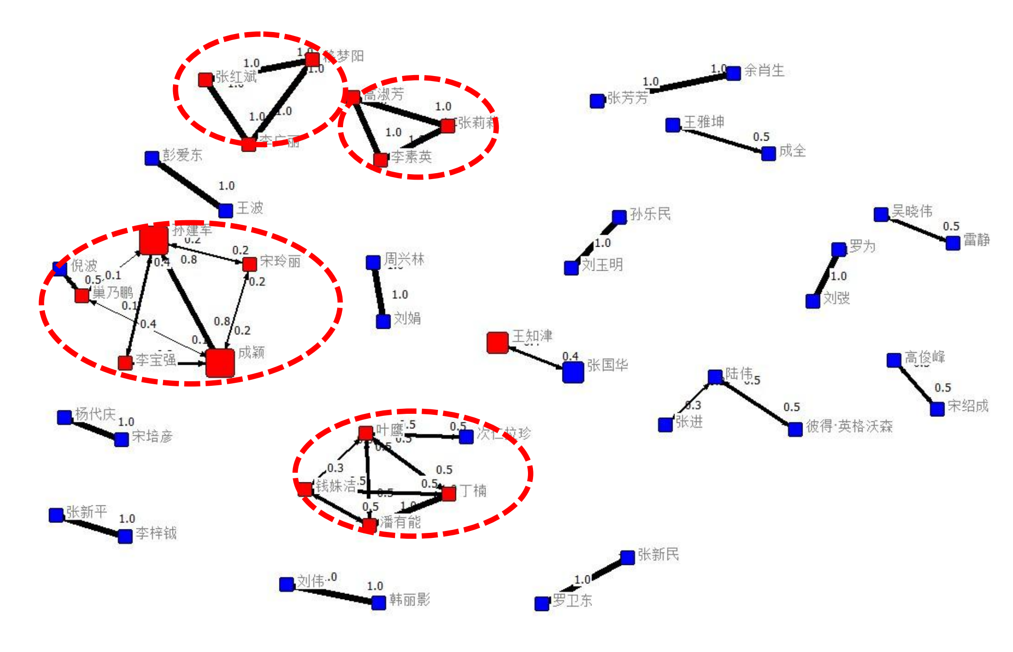


图 3 高频作者合著网络图谱

利用 Ucinet 的中心性测度功能对作者合著网络图谱进行分析，取度数中心度大于 2 的作者，共有 15 位作者，如表 4 所示。由度数中心度（degree）属性可知，取值最高的是成颖、孙建军和叶鹰，度数中心度值为 4，表示这三位作者曾与 4 位作者共同合作撰写情报学相关性领域的论文，在合作发文方面较为活跃；而中介中心度（betweenness）取值最高的是巢乃鹏，取值为 4，其次是叶鹰、成颖、孙建军，取值分别为 3、2.5、2.5，并且与其他作者的中介中心度取值相比高出较多，这表明在该合著网络结构中，上述学者具有较大的影响力，掌握着该领域较多的资源；由该表的 K 核属性（K-core）可知，在情报学相关性领域高频作者的合作情况中，最多只有 4 位作者共同撰写论文，这表明情报学相关性领域合作撰文的较少；根据接近中心度（closeness）属性取值显示，各位学者与其他学者的平均最短距离值均在 1800~2000 左右，数值较大，表示各个作者之间的联系程度非常弱，再次印证学者之间缺乏学术交流与合作；最后，关于网络密度（density），根据 Ucinet 计算得出的结果，高频作者合著网络的整体平均网络密度是 0.0049，这个数值所代表的网络紧密程度是非常低的，再次表明 IC 领域的作者合作度较低。

chinaXiv:201705.00900v1

ID	*K-core	Degree	Betweenness	Closeness	Density
成颖	2	4.000	2.500	1892.000	0.500
孙建军	2	4.000	2.500	1892.000	0.500
叶鹰	3	4.000	3.000	1936.000	0.500
巢乃鹏	2	3.000	4.000	1893.000	0.333
潘有能	3	3.000	0.000	1937.000	1.000
丁楠	3	3.000	0.000	1937.000	1.000
钱姝洁	3	3.000	0.000	1937.000	1.000
李宝强	2	2.000	0.000	1895.000	1.000
宋玲丽	2	2.000	0.000	1895.000	1.000
李广丽	2	2.000	0.000	2026.000	1.000
李素英	2	2.000	0.000	2026.000	1.000
张莉莉	2	2.000	0.000	2026.000	1.000
移梦阳	2	2.000	0.000	2026.000	1.000
张红斌	2	2.000	0.000	2026.000	1.000
高淑芳	2	2.000	0.000	2026.000	1.000

表 4 高频作者合著中心性测度

通过 Ucinet 的 K 从（K-plex）功能，可以分析出某领域作者合作的小团体，借助这一功能，对 IC 领域的作者进行了 K 从分析。需要说明的是，在设定 K 值和最小的小团体节点数时，要根据实际情况进行反复调整。K 值越小，最小的小团体节点数越大，则条件越苛刻，则小团体越少[7]。经过调整，最终设定 K 值为 2，最小的小团体节点数为 3，得到 12 组团体，如下图 4 所示。通过 K 从分析的结果可以得到图 5 作者聚类图，由图可知共有 5 个作者小团体。其中图 5 为作者网络的整体聚类情况，图 5-1 至图 5-4 是形成的作者小团体的聚类图，由图 5 中截取获得。将由图可知，成颖、孙建军、巢乃鹏、李宝强、宋玲丽、倪波是一个小团体，叶鹰、次仁拉珍、潘有能、丁楠、钱姝洁是一个小团体，张进、陆伟、彼得•英格沃森是一个小团体，李广丽、移梦阳、张红斌是一个小团体，李素英、张莉莉、高淑芳是一个小团体。这一结论也可由图 3 作者合著网络图谱中得到验证。另外，从表 5 高频作者发文量统计表中可以看出，在情报学相关性领域，成颖发文量最多，共 6 篇论文，其次是王知津和孙建军，均为 5 篇。

chinaXiv:201705.00900v1

```
12 k-plexes found.

   1:  成颖 孙建军 巢乃鹏 李宝强
   2:  成颖 孙建军 巢乃鹏 宋玲丽
   3:  成颖 孙建军 李宝强 宋玲丽
   4:  成颖 巢乃鹏 倪波
   5:  孙建军 巢乃鹏 倪波
   6:  叶鹰 次仁拉珍 潘有能
   7:  叶鹰 次仁拉珍 丁楠
   8:  叶鹰 次仁拉珍 钱妹洁
   9:  叶鹰 潘有能 丁楠 钱妹洁
  10:  张进 陆伟 彼得·英格沃森
  11:  李广丽 移梦阳 张红斌
  12:  李素英 张莉莉 高淑芳
```

图 4 K 从计算结果

chinaXiv:201705.00900v1

3.000 1.600 1.333 1.133 1.057 1.000 0.600 0.000

王知津 3
乔建忠 4
黄名选 6
张新民 7
罗为 8
刘辉 9
马志杰 10
张云秋 11
刘骏 12
曹红兵 13
周军 14
赵印红 15
武夷山 17
苏君华 18
李培 19
吴晓伟 21
成全 26
屈鹏 27
罗卫东 28
张国华 30
高俊峰 31
常娥 32
韩丽影 33
孙乐民 34
康永林 35
张晓林 36
杨秀丹 37
赵小康 38
李梓铖 39
余肖生 40
朱俊卿 42
雷雪 43
张俊丽 45
温有奎 46
周琪 47
徐斌 48
罗云锋 49
冯勤超 50
唐小莹 51
王冰 52
张进 25
陆伟 29
彼得·英格沃森 53
杨代庆 54
李爱红 56
张朋柱 57
王建芳 58
刘玉明 59
王雅坤 60
周兴林 61
周婴 62
陆嘉德 63
叶鹰 5
次仁拉珍 22
潘有能 23
丁楠 24
钱妹洁 64
刘海峰 65
彭爱东 66
王继民 67
张芳芳 68
王益明 69
伍宾 70
刘琪 72
李静 73
张长江 74
陈陶 75
田俊华 76
刘博晓 77
宋绍成 79
宋培彦 80
马小闯 81
刘伟 82
雷静 83
郭晓苗 84
孔令夷 85
陈明英 86
牛昱昕 87
沈迪飞 88
濮德敏 89
李广丽 44
穆梦阳 78
张红斌 90
李素英 55
张莉莉 71
高淑芳 91
黄美君 92
王波 93
熊回香 94
邹佳骏 95
贾朝光 96
黄晨 97
李宝强 20
宋玲丽 41
成颖 1
孙建军 2
单乃鹏 16
倪波 98
刘娟 99
张新平 100

图 5 作者聚类图

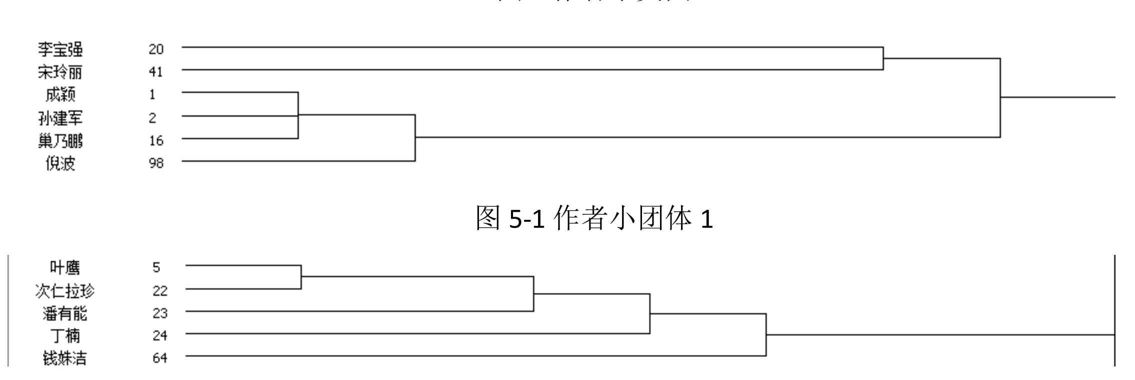


图 5-1 作者小团体 1

chinaXiv:201705.00900v1

图 5-2 作者小团体 2

张进 25
陆伟 29
彼得·英格沃森 53

图 5-3 作者小团体 3

李广丽 44
移梦阳 78
张红斌 90

图 5-4 作者小团体 4

李素英 55
张莉莉 71
高淑芳 91

图 5-5 作者小团体 5

字段: 作者	记录数	% (337)	图形
成颖	6	1.780 %	
孙建军	5	1.484 %	
王知津	5	1.484 %	
乔建忠	4	1.187 %	
叶鹰	4	1.187 %	
黄名选	3	0.890 %	
丁楠	2	0.593 %	
刘弢	2	0.593 %	
刘辉	2	0.593 %	
吴晓伟	2	0.593 %	
周军	2	0.593 %	
屈鹏	2	0.593 %	
巢乃鹏	2	0.593 %	
张云秋	2	0.593 %	
张国华	2	0.593 %	
张新民	2	0.593 %	
张进	2	0.593 %	
成全	2	0.593 %	
曹红兵	2	0.593 %	
李培	2	0.593 %	
李宝强	2	0.593 %	
次仁拉珍	2	0.593 %	
武夷山	2	0.593 %	
潘有能	2	0.593 %	
罗为	2	0.593 %	
罗卫东	2	0.593 %	
苏君华	2	0.593 %	
赵迎红	2	0.593 %	
陆伟	2	0.593 %	
马志杰	2	0.593 %	
高俊峰	2	0.593 %	

表 5 高频作者发文量统计表

综合上述指标，可以得出如下结论。

首先，研究情报学相关性的学者数量较少，论文成果数量有限，如上表 4 所示，共有 337 名学者发文，共有 183 篇论文成果，在该领域发表论文数量最多的学者共发表 6 篇，且这 6 篇的被引次数不理想，最高被引 12 次，有 2 篇没有被引用。这一方面表明图情领域学者对该相关性领域的关注度不足，重视程度不

够；另一方面表明国内缺乏代表性学者，这一点与国外有较大差距。

其次，对学者所在机构进行统计，发现发文量最多的为中国科学院国家科学图书馆（中国科学院文献情报中心）和南京大学，均发文 12 篇，其次分别为武汉大学（11 篇）、中国科学技术信息研究所（7 篇）、南开大学（7 篇）等。发文机构信息统计表明，对情报学相关性研究给予较大关注的是图情研究院所及和重点高校信息管理学院，这表明图情机构对情报学相关性研究具有重要作用，要充分发挥其自身优势，推动该领域的发展。

最后，根据 K 核、接近中心度以及网络密度的取值可知，情报学相关性领域的作者合著网络整体较为松散，作者之间缺乏交流与合作。将所有参与合著的作者进行统计，发现只有 48 位作者参与过合著，其中共形成 4 个小团体。这表明在该领域的研究较为分散和孤立。但是对于机构之间的合作，有 20 篇文章为跨机构合作，合著较为活跃的地区为北京、上海、江苏（南京、苏州）、湖北（武汉）、广东（广州）等地，这表明发达地区对该领域的研究较为活跃。

4. 小结

本文通过运用社会网络分析方法，从关键词共现和作者合作两个角度对情报学相关性的研究现状进行分析。

通过关键词共现网络分析，对目前情报学相关性领域的研究热点予以阐述。本文认为，国内对情报学相关性的研究重点在于信息检索的相关性研究、引文分析中的相关性研究、数据挖掘与知识管理中的相关性研究以及与国外情报学相关性研究的对比研究。总体看来，国内对情报学相关性的研究更为偏重实际应用，尤其是对检索领域的相关性研究。由于网络化的发展，搜索引擎检索的相关性已成为重要的研究领域，尤其是对搜索算法改进、检索系统性能评价等方面。但是另一方面，也反映出国内对相关性的研究领域较为局限，一方面是理论研究不足，对相关性概念的确定、相关性模型等方面的探讨较少；另一方面是相关性的应用研究过于局限于检索的相关性，根据统计，在所有发表的相关性论文中有多于一半的论文都是研究检索的相关性，这反映出研究点较为单一。因此，为了完善情报学相关性的研究与发展，应该从以下两方面努力，一是注重该领域的理论研究，尤其应借鉴国外丰富而完善的理论研究成果；二是注重将相关性应用于其他领域，如引文分析、数据挖掘、知识地图等方面。

chinaXiv:201705.00900v1

通过高频作者合作网络分析，本文认为，目前情报学相关性领域的不足之处体现在两方面，一方面是缺乏研究动力，发文作者之间学术交流较少，形成的学术小团体较少，合作地区有限，彼此间的研究较为封闭和孤立，导致该领域的理论研究缺乏增长动力；另一方面，情报学相关性领域的研究学者中缺乏权威性、引领性的核心研究者，这表明学者对该领域的研究积极性不高，对该领域的重视程度不足，不利于情报学相关性领域未来的发展。为此，应从以下几方面予以改进。首先，可在全国范围内开展以“情报学相关性研究”为主题的学术研讨会，邀请国内外该领域的著名学者，共同交流该领域重要的理论与方法，使得该领域的概念及理论得以普及，同时吸引更多的学者关注该领域，从而激发其研究动力。其次，鼓励学者积极与国外该领域的相关专家学者交流学习，借鉴和吸收国外情报学相关性的相关理论，为我国在该领域的研究注入新的活力。最后，尽管图情领域是相关性研究的重要力量的，但是由于相关性涉及到相关的搜索算法、搜索引擎等计算机学科的内容，因此应加强与计算机等其他学科相关学者的交流与学习。此外，由于相关性的研究还涉及到搜索引擎等实际应用问题，因此除了学界做出努力之外，还应积极发挥业界的力量，加强与搜索引擎公司的交流，以便在实践中获得理论之源，发现新的研究点，从而促进我国情报学相关性研究的蓬勃发展。

chinaXiv:201705.00900v1

参考文献：

[1]Saracevic. Relevance: A review of and a framework for the thinking on the notion in information science. Journal of the American Society for Information Science, 1975,26:321–343.

[2]Stefano Mizzaro. Relevance: The whole history[J].Journal of the American Society for Information Science,1997,48(9):810-832

[3]庞弘燊,徐文贤. 近年来国外信息检索的相关性研究进展[J]. 中国图书馆学报,2009,04:88-94.

[4]邱均平,伍超. 基于社会网络分析的国内计量学作者合作关系研究[J]. 图书情报知识,2011,06:12-17.

[5]赵蓉英,李飞. 基于社会网络分析方法的国内外信息计量比较研究[J]. 情报科学,2013,02:7-12.

[6]郑艳超. 基于数据挖掘的竞争情报与知识管理融合模型研究[J]. 图书馆学研究,2010,17:28-31.

[7]徐媛媛,朱庆华. 社会网络分析法在引文分析中的实证研究[J]. 情报理论与实践,2008,02:184-188.